Study of Image Tracking for Active Vision Systems(能動視覚システムのための画像追跡法の研究)

著者	佐藤 美徳
号	315
発行年	2004
URL	http://hdl.handle.net/10097/13015

氏名（本籍）	佐藤（さとう） 美徳（よしのり）（岡山県）
学位の種類	博士（情報科学）
学位記番号	情博第315号
学位授与年月日	平成17年3月25日
学位授与の要件	学位規則第4条第1項該当
研究科、専攻	東北大学大学院情報科学研究科（博士課程）システム情報科学専攻
学位論文題目	Study of Image Tracking for Active Vision Systems（能動視覚システムのための画像追跡法の研究）
論文審査委員	（主査）東北大学教授 出口光一郎
	東北大学教授 中野 栄二　東北大学教授 橋本 浩一
	東北大学教授 青木 孝文　東北大学助教授 岡谷 貴之

論文内容要旨

1 Introduction

The goal of this thesis is to develop a method for tracking automatically a moving object in an image sequence, and it is one of the most important problems in computer vision. We call this problem "*image tracking.*" This problem has been studied by many researchers, and many algorithms have been proposed for applications in various fields, such as surveillance systems, perceptual interfaces, video cording, robotics and amusements.

The concept of "*active vision*" is to model the interaction between a target object, sensors (cameras) and their surrounding environment and to plan the action of cameras for a certain task, such as surveillance, robot manipulation and recognition of surrounding environment. This concept has also been applied to image tracking problems. For example, in construction of surveillance systems, planning the camera movement and controlling its pan, tilt and zoom are required for finding out target persons and tracking them. In this thesis, we consider important aspects for building image tracking systems by using active vision cameras.

In image tracking problems, we must often cope with misleading information that is caused from cluttered background and partial occlusion. In order to overcome the misleading information, this thesis especially consider the following aspects:

- feature representation of target objects, and
- prior knowledge about the targets, such as their appearance changes and their motions.

The robustness and efficiency of image tracking rely on how these aspects are combined.

While a target object moves in space, we usually have to deal with its appearance changes. When it is observed by active vision cameras, it is difficult to use information of the background because of their own movements. Therefore, we need to come up with better representation for capturing the images that can handle its appearance changes.

If image sequences are given in advance, we can obtain the dynamics of the target object, such as the target motion and its appearance changes, from their temporal relation. For obtaining the dynamics, we usually need learning. Here, we treat the obtained dynamics as "*prior knowledge.* The dynamics is often described as a dynamical model in statistical methods such as filtering. From the dynamical model, we can estimate the target motion in the tracking procedure based on statistical methods.

In this thesis, with these aspects in mind, we investigate tracking algorithms in two approaches: boundary-based tracking and region-based tracking.

2 Boundary-based tracking using prior knowledge of shape changes

In order to track a moving object and know its pose, this chapter introduces the boundary-based tracking approach. This tracking approach relies on the information about the object boundaries, and has often employed active contouring. The active contouring includes tasks to detect the target contour and track it.. If the accurate contour of the target is obtained, it gives us the information to know its pose. However, when the target moves in an image sequence, we often suffer from cluttered background or partial occlusion and it is difficult to extract the target contour. Here, we construct an algorithm of active contouring for extracting the accurate contour.

In earlier study, the energy-based active contouring methods have been introduced for extracting the target contour from an image sequence. In this framework, the desired contour is estimated by minimizing these energy functions according to the change of the energy function itself. For example, Cremers et al. (2003) also introduced an energy-based active contouring method and succeeded in stably extracting the contour of a target object even if there was misleading information due to noise, clutter and occlusion in input images. In their method, they modified the Mumford-Shah segmentation method for the external energy and integrated prior knowledge of the target shapes into it. However, their method was effective only for static images, and depended only on each image regardless of the temporal association between images even when image sequences were given for the training data.

In this chapter, in order to track the contour of a moving object smoothly in an image sequence, we extend the method proposed by Cremers et al. basically and introduce the new energy function to deal with temporal appearance changes and motion of the target object. Here, we consider two kinds of the prior knowledge and construct an algorithm based on the energy functions including these. The first prior knowledge is about the target contours obtained from various viewing points in advance. This prior knowledge enables stable extraction of the target contour even when the target object has misleading information which is caused from cluttered background and partial occlusion. The second knowledge is about the shape changes of the target. This prior knowledge is prepared by analyzing the relation between images from image sequences given in advance, and enables smooth tracking even when the appearance of the target changes frequently in their views because of the movements of cameras.

In the end of this chapter, we show experimental results to confirm effectiveness of this method. In this experimental, we attempted to track a moving and rotating object. The proposed method could extract and track the target contour in spite of the cluttered background and the occluded part.

3 Color-based tracking by using particle filter

In order to track a non-rigid object, the color-based tracking approach is discussed from Chapter 3 to Chapter5. The color-based approach has many advantage, which are its rotation and scale invariance and the robustness to motion blur. In this framework, the representation that we use treats the target as the color histogram of the region where it exists in an image. This thesis introduces the color-based methods based on the particle filter especially.

In the conventional tracking algorithms using the particle filter such as the Condensation algorithm, the latest observation obtained from the current frame in an image sequence is used only in the weighting step and not in the sampling step. As a result, a large number of particles are often required for approximating the posterior probability density of the state of a target object properly.

In this chapter, we apply the Kalman particle filter (KPF) to the color-based tracking approach in order to cope with the misleading information. This KPF is one of the particle filters and includes the principle of the updating rule of the Kalman filter. We modified this KPF for color-based tracking. This modified KPF can properly approximate the probability density of the position of a target object by using the latest observation, and needs fewer particles for image tracking than conventional approaches using particle filters.

In some experiments, the performance of the proposed KPF tracking method was compared with that of the Condensation tracking method, one of the conventional methods. Our proposed method required smaller number of particles for tracking than the Condensation method. However, since the computational cost per particle is larger in the KPF tracking method, we cannot easily imply that the KPF tracking method has the advantage of the computational efficiency. Therefore, we employed the "critical size" and compared these algorithms by using particles whose number is the critical size. As the result of the comparison, we could confirm that the KPF tracking has the advantage of computational efficiency. So that, the computational cost for color-based tracking can be reduced by using the KPF.

4 Image tracking by using graphical models

When the partial occlusion of the target occurs, the estimated trajectory of the color-based method using particle filter is not stably obtained. This is because of evaluating the whole target as one region. In other words, the color histogram of the target was made from one region including the target in the image. If the target can be divided into some parts (e.g. the head and the body of a person), it should be modeled as a multiple region. In this chapter, we propose to treat the multiple region as a graphical model. In the use of this graphical model, we define "*messages*." The messages include the state of a part of the target and are sent to other parts. As a result, if the target object has occluded part, the lack of the information of the occluded part can be compensated by the messages of the visible parts, and the trajectory of the target can be obtained stably.

In this framework, the estimation at each frame was obtained from temporal and spatial information. The temporal information was given by the conventional particle filter, and the spatial information was given by the message propagation in the graph. Because the estimation given by the particle filter was updated by the additional operation, the message propagation, the problem of the conventional approaches, which is that they cannot effectively use information of observations in particle filtering, went away in hindsight.

In the end of this chapter, we conducted to track a pedestrian in image sequences to confirm the effectiveness of our methods. In these experiments, we modeled the pedestrian as a graph. As the result of our proposed method, the position of the person was estimated stably even when the part of the target was occluded partially, and the position of the occluded part could be estimated.

5 Image tracking by using dispersed trackers

When two or more candidates of a target object exist in an image, for example, when there are some persons who wear similar clothes in an image, we tend to miss the target. In order not to fail image tracking, we propose a method to track all candidates similar to the target in parallel when the observation is unstable.

The particle filter can approximate the probability density of a target state according to the weighted particles. In the traditional tracking methods using the particle filter, each particle dispersedly observes the target. However, the success and accuracy of those tracking methods considerably depend on the initialization of the parameters in the particle filter even when images had misleading information. As a result, we take a lot of works for initializing dynamical models and parameters of the particle filter.

If we change the value of the variance of the noise that is added in the prediction step of the particle filter, we obtain different tracking results. Here, we especially focus on the variance of the noise and propose the dispersed tracker based on this property. This tracker has some particle-sets and each particle-set of the tracker is sampled according to the different noise variances. When each particle-set obtains different estimation, we can find that some candidates exist. The number of the dispersed tracker increases or decreases according to the estimation, and the trackers can track two or more candidates of target dispersedly and observe these in parallel.

In the end of this chapter, we conducted some experimental results to confirm the effectiveness of our methods. In these experiment, we tried to track a target with dispersed trackers in image sequences that have some candidates similar to the target. As the result of these experiments, the dispersed tracker did not miss the target and could track both the target and the candidates.

6 Conclusions

This chapter concludes the achievements of this thesis. This thesis aimed to construct image tracking algorithms for building active vision systems. Toward the aim, we noted the robustness of image tracking relies on the target representation and the use of the prior knowledge about the target, and addressed some problems of image tracking in two categories: the boundary-based tracking and the region-based tracking. In order to overcome them, we constructed some algorithms and could obtain some effective results.

論文審査の結果の要旨

動画像中の移動体を自動的に追跡する技術は、例えば、セキュリティーシステム、知覚インターフェイス、ロボティクス等の様々な分野から要求されている。さらに近年では、人間の目のように自ら動きまわることが出来るカメラを用いる視覚システムである能動視覚の実現が多くの分野で望まれており、画像追跡はその鍵となる技術である。本論文では、高度な知能を持つ能動視覚システムを構築するために、能動視覚特有の困難さを明確にし、新たな移動体追跡法の提案を行い、実験によりその有効性の確認を行っている。本論文はその成果をまとめたもので、全文6章よりなる。

第1章は序論である。

第2章では、移動対象（剛体）の姿勢を知る必要がある場合を想定し、画像輪郭に基づく追跡法について論じている。対象の姿勢に関する知識として、事前に学習した追跡対象の輪郭およびその見え方の変化に関する知識を、エネルギー関数に基づく輪郭抽出法に組み込むことを提案している。その結果、紛らわしい情報が画像中に混入した場合（背景が複雑である、または対象の部分が隠れる等）でも安定に輪郭を抽出し、その輪郭を追跡することに成功した。そして、実験によりその有効性が実証されている。

第3章から第5章では、非剛体の対象でも追跡できる方法として、逐次モンテカルロ法を用いた色分布に基づく追跡法について論じている。その中で、第3章では、従来の逐次モンテカルロ法を用いた追跡法における問題点の考察を行っている。特に画像から得られる色分布情報を効果的に使用するために、逐次モンテカルロ法の一つである"Kalman particle filter"を修正し、それを色分布に基づく追跡法の枠組みに適用した。その結果、同等の計算コストであれば、従来の追跡法よりも精度の高い位置推定を行うことに成功しており、その有用性を示している。

第4章では、色分布に基づく移動体追跡法に、グラフィカルモデルを導入した方法について論じている。対象を分割して知覚できる場合、その対象全体を一つのグラフとして扱い、そのグラフ構造を逐次モンテカルロ法に組み込んだ追跡手法を提案している。この結果、時空間のデータの繋がりを追跡に生かすことができ、実時間で移動物体を追跡する場合においても対象の軌跡をより正確に求めることに成功している。そして、その有効性も実験により実証している。

第5章では、分散して対象を追跡する方法について論じている。移動体追跡中、対象と似た物体が混入した場合、追跡対象を間違えてしまう状況が存在する。これを回避するために、対象である可能性を持つもの全てを分散して追跡する方法を提案し、それを逐次モンテカルロ法の原理に基づき実現している。その結果、効率的に分散して対象を追跡することに成功している。これは応用上重要な成果である。

第6章は結論である。

以上要するに本論文は、能動視覚システムの構築のために、新たな移動体追跡法の提案を行い、その有効性を示したものであり、コンピュータビジョンおよび情報科学の発展に寄与するところが少なくない。

よって本論文は博士（情報科学）の学位論文として合格と認める。